Laboratories for Population Statistics
Reprint Series No. 3
December 1971

THE DESIGN OF AN EXPERIMENTAL PROCEDURE FOR OBTAINING ACCURATE VITAL STATISTICS

SOME RESULTS FROM ASIAN POPULATION GROWTH STUDIES

10985

10985

El Programa Internacional de los Laboratorios Estadisticos de la Población, se dedica a la investigación de mejores métodos para la medida de los cambios ocurridos en la población y al análisis de datos demográficos. Para este fin, el Programa espera establecer y mantener "laboratorios" experimentales en los cuales nueva metodología pueda ser comprobada y donde aplicable técnicas innovadas para et análisis de datos ya obtenidos puedan ser desarrolladas. Los proyectos se basan en acuerdos cooperativos entre la Universidad de Carolina del Norte, en Chapel Hill, é instituciones científicas de varios países. El Programa se lleva a cabo gracias a la participación de estas instituciones y a la asistencia de la Agencia para el Desarrollo Internacional del Gobierno de los Estados Unidos. Publicaciones científicas y técnicas son distribuídas, sin costo alguno, a un número limitado de instituciones y personas interesadas en problemas relacionados con medidas demográficas de la población. Las diferentes opiniones y conclusiones pertenecen a los autores respectivos, por lo que no deben ser atribuídas a las instituciones patrocinadoras.

Le Programme international des laboratoires statistiques de population s'adresse au perfectionnement des procédés qui mesurent les variations dans la population et à l'analyse des données démographiques. C'est avec ce but que le programme cherche à établir et à maintenir des "laboratoires" experimentals où de nouvelles métodes démographiques peuvent être verifiées et où applicable, des voies innovatrices d'analyse des données existantes seront développées. Les projets sont basés sur des accords cooperatifs entre l'Université de la Caroline du Nord à Chapel Hill et des institutions scientifiques en divers pays. Ce programme se réalise grâce à la participation de ces institutions et l'assistance de l'Agence pour le Développement International du Gouvernement des Etats-Unis. Les publications scientifiques et techniques sont distribuées gratuitement aux institutions et aux individus qui traitent des problèmes démographiques. Les opinions et les conclusions exprimées sont celles des auteurs et non celles des institutions nommées ci-dessus.

The International Program of Labora
cated to research on improved metho
change and to the analysis of demog
gram seeks to establish and support e
measurement methodology can be tes
sis of existing data can be develope
agreements between the University
scientific institutions in various coun
through the participation of these i
Agency for International Developme
States. Scientific and technical issu
number of institutions concerned wi
The opinions and conclusions expres
and should not be attributed to the s

Community Health Cell
Library and Information Centre
359, "Srinivasa Nilaya"
Jakkasandra 1st Main,
1st Block, Koramangala,
BANGALORE - 560 034.
Ph : 2553 15 18 / 2552 5372
e-mail : chc@sochara.org

Laboratories for Population Statistics
Reprint Series No. 3
December 1971

THE DESIGN OF AN EXPERIMENTAL PROCEDURE FOR OBTAINING ACCURATE VITAL STATISTICS

Ansley J. Coale

SOME RESULTS FROM ASIAN POPULATION GROWTH STUDIES

William Seltzer

International Program of Laboratories for Population Statistics
Forrest E. Linder, Director

The Department of Biostatistics
School of Public Health

The Carolina Population Center

The University of North Carolina
Chapel Hill, N.C. 27514
U.S.A.

Il est pratiquement indispensable de disposer de registres des naissances et des décès complets et exacts pour traiter correctement les problèmes administratifs et juridiques et pour assurer le succès des programmes sociaux et économiques, notamment dans les domaines de la santé et de la médecine.

Dans les régions économiquement moins développées où le besoin de données exactes concernant les naissances et les décès est le plus vif, l'enregistrement des déclarations d'état civil est souvent inexistant, peu sûr, ou fort incomplet. De nombreuses années sont nécessaires pour ètablir un registre complet. Dans l'intervalle, l'enregistrement dans des zones d'échantillonage spécialement choisies peut fournir des estimations sur les taux nationaux de naissances et de décès et constitue une expérience pilote dans l'établissement d'un registre complet. Il est de la plus haute importance de connaître le degré d'intégralité d'un enregistrement dans la zone d'échantillonage. Il est proposé dans cet article que la portée des résultats soit vérifiée au moyen de la méthode bien connue du couplage individuel des événements enregistrés avec un registre indépendant des mêmes événements. Le registre indépendant des naissances et des décès serait confectionné à partir d'enquêtes des ménages. Le recours au couplage individuel permet d'évaluer l'intégralité tant du registre que des événements enregistrés lors de l'enquête des ménages.

La caractéristique d'évaluation qui ressort du système proposé pourrait contribuer à la mise au point d'un enregistrement et à la conduite d'enquêtes plus efficaces.

A complete and accurate register of births and deaths is virtually indispensable for the proper handling of administrative and legal problems, and for the success of social and economic programs, especially in health and medicine.

In economically underdeveloped areas, where the need for accurate data regarding births and deaths is most acute, vital registration is frequently non-existent, unreliable, or far from complete. It takes many years to establish a complete register. Meanwhile, registration in specially chosen sample areas can provide estimates of national birth and death rates, and constitute pilot experience in the establishment of a complete register. It is of crucial importance that the degree of completeness of registration in the sample area be known. It is proposed here that coverage of the results be ascertained by the well-known technique of individual matching of registered events with an independent record of the same events. The independent record of births and deaths would be assembled through household interviews. The use of individual matching makes it possible to evaluate the completeness both of the register and of the events recorded in the household survey.

This evaluation feature of the proposed scheme could help in the design of more effective registration and more effective surveys as well.

Un registro de nacimientos y defunciones completo y exacto es virtualmente indispensable para el trato apropiado de problemas administrativos y jurídicos, tanto como para el suceso de programas sociales y económicos, especialmente de la salud y medicina.

En áreas económicamente subdesarrolladas, donde existe una necesidad aguda de datos de defunciones y nacimientos, el registro vital es frequentemente no-existente, no confiable, o lejos de ser completo. Establecer un registro completo lleva muchos años. Mientras tanto, el registro de datos en áreas especialmente seleccionadas puede proporcionar estimadores de tasas de nacimientos y defunciones, y constituir una experiencia piloto en el establecimiento de un registro completo. Es de importancia crucial conocer el grado de integridad del registro en el área de la muestra. En este trabajo se propone indagar la cobertura de los resultados por medio de la técnica ya bien conocida de confrontación individual de hechos registrados, con un registro independiente de esos mismos hechos. El registro independiente de nacimientos y defunciones se obtendría por medio de entrevistas de hogares. El uso de la confrontación individual hace posible evaluar la integridad del registro y de los hechos recogidos en la encuesta de hogares.

La característica de evaluación del esquema presentado, podría ayudar a diseñar un registro más efectivo así como encuestas más efectivas.

Cet article présente et analyse les renseignements estimatifs touchant aux naissances et aux décès, dont la portée a été établie en comparant les divers éléments d'une série relativement complète d'études réalisées en Asie en utilisant les systèmes à double collecte des événements vitaux. Voici les principales conclusions fondées sur ces études: (1) les renseignements sont en général plus complets dans le cas des naissances que dans celui des décès; (2) en dépit de cela, on relève une tendance vers une sous-estimation des taux non ajustés de l'accroisement naturel; (3) on note, dans chaque type de système de collecte des données, une grande variabilité en ce qui concerne la portée des taux résultant des observations; (4) dans l'ensemble, il semble que les enquêtes des ménages, lorsqu'elles ne sont pas répétées, donnet lieu à des comptes rendus incomplets.

Coverage estimates of births and deaths derived from the matching experience of a relatively comprehensive group of studies conducted in Asia and employing dual collection systems are presented and analysed. Major conclusions based on these Asian studies include: (1) coverage of live births has generally been more complete than that for deaths; (2) nevertheless, unadjusted rates of natural increase tend to be underestimates; (3) there is a great variability within each type of collection system in the coverage rates observed; and (4) one-time household surveys, as a group, seem to be particularly subject to underreporting of vital events.

Se presentan y analizan las estimaciones sobre la cobertura de nacimientos y defunciones derivadas de la experiencia de confrontar un grupo de estudios, relativamente comprehensivo, conducidos en Asia, los que han sido hechos empleando un sistema de recolección dual. Las conclusiones principales de estos estudios asiáticos son: 1) la cobertura de nacimientos vivos ha sido generalmente más completa que la de las defunciones; 2) no obstante, las tasas sin ajustar de crecimiento natural tienden a estar subestimadas; 3) hay una gran variabilidad en la cobertura de las tasas observadas dentro de cada tipo de sistema de recolección; y 4) encuestas de hogares de única-vez como grupo, parecen ser incompletas en cuanto a la información de hechos vitales.

THE DESIGN OF AN EXPERIMENTAL PROCEDURE FOR OBTAINING ACCURATE VITAL STATISTICS

Ansley J. Coale

Probably a majority of the world's population lives in areas where the registration of births and deaths is either absent or so incomplete as to be an unreliable indicator of birth and death rates. In the long run there is no substitute for a complete register of vital events. The registration of individual births has proven an indispensable adjunct to resolving legal and administrative problems, and the registration of deaths (including an accurate recording of the cause of death) is invaluable for the administration of medical and health programs.

It will be years, however, before many of the countries just embarking on a program of industrialization and modernization will be able to attain complete or nearly complete registration of all births and deaths. In the meantime, accurate knowledge of current levels of fertility and mortality and of differential rates according to various characteristics of the population are indispensable to an understanding of the population problems in each country, to the formulation of policies with respect to family limitation, and to the design and operation of an efficient public health program. In the interval, while a complete registration system is being developed, and, indeed, as an important aid in developing such a program, the recording of births and deaths for a representative sample of the national population would be of the highest value.

The following notes represent an attempt to outline in specific detail a proposal for obtaining accurate data on the number of births and deaths in a sample area. The author can make no claim of originality in setting forth this proposal. Use was made of individual matching of data from a register and from a survey in the United States Censuses of 1940 and 1950, and the technique was refined and tested with data from the Singur Health Unit near Calcutta by Chandrasekar and Deming.* The only novelty in the scheme proposed here is the application to a sample area. The idea of sample registration areas (without the proposed system of verification) has been widely entertained.

The underlying technique is to obtain data on individual events in the sample area by two independent means. The events recorded by each means of measurement must be identified accurately enough to establish a "match" of the two records of the same event. The two systems of measurement would be as complete a register as possible of births and deaths in the sample area, and some form of periodic household survey--probably every year--that asks for births and deaths during the interval since the preceding survey.

---

*Chandrasekar, C. and Deming, W. Edwards, "On a method for estimating birth and death rates and the extent of registration," Journal of the American Statistical Association, Vol. 44 (1949), pp. 101-115. [Reprinted as No. 1 in the Reprint Series of the Laboratories for Population Statistics (May 1971)]

Reprinted from *International Population Conference, New York, 1961*, Vol. 2, London, International Union for the Scientific Study of Population, pp. 372-376 with permission of the publisher.

The simplest form of experiment would require the selection of a scientific sample of villages in the region where the experiment is to be conducted, and the initiation of the two systems of data collection in all these villages. The registration might be either some experimental form of improvement of an existing registration system, or alternatively a wholly new registration scheme, employing, for example, health workers or school teachers. A possible form of improvement of existing machinery might be renumeration of the local registrar on a piece-rate basis, with some offsetting penalty for false registration. The household survey could experiment with various probing devices, such as asking for pregnancies and subsequently inquiring how each pregnancy has terminated (and similarly with illnesses); asking for children born who are still alive and then about children who have died, etc.

The most difficult feature of an experiment of this sort will be to establish the match between events recorded in the periodic household survey, while at the same time maintaining the independence of the two sources of information. Experience with matching shows surprisingly frequent instances of uncertainty as to whether the two events are in fact different or the same. The address reported for a birth in the register and a birth in the household survey may be the same, but the name or perhaps the sex recorded for the child may be different. The problem is complicated by the fact that it is not only necessary to establish identity or lack of identity, but to find out whether an event reported by one system and not the other should indeed be included in the universe of properly recorded vital events. For example, if a birth is reported in a household survey and not in the register there is always the possibility that the birth may have occurred in the preceding year on the one hand, or in another village on the other. In either case it would be inappropriate to count this as an omission from the register. In practice it may be necessary to effect a reconciliation in each instance where one system records an event not recorded by the other system. The problem of independence will doubtless become increasingly acute as an attempt is made to continue the two systems side by side in the same village. It will be very difficult to convey the nature of the scientific problem involved to the local registrar and to the surveyors who conduct the household interview. They will doubtless learn, because of the reconciliation procedures, that the results of each person's work are being used to check the other's. There will be a strong tendency for the two lists to approach identity through collaboration rather than completeness. It will take good administrative machinery and very close supervision to see that such collaboration does not occur.

In spite of the major difficulties just described there seems little doubt that an event-by-event matching of two independent systems represents the most promising technique for assessing the accuracy of any system of recording. A comparison without individual matching between two systems serves only to compare their relative completeness. A register and a survey might, for example, both indicate a birth rate of 38 and a death rate of 18 for some Indian state, when the true figures were 45 and 25 respectively. Agreement with regard to number of events between register and survey, however, would indicate no more than that they both achieve about the same degree of completeness. In contrast, the individual matching of records makes

it possible to estimate the completeness of both systems. In fact, the only major deficiency of the matching procedure (assuming that the problem of matching itself can be adequately solved) is the existence of a correlation between the probability of omission from each of the two systems. The existence of correlation makes it difficult to estimate the number of instances of events undetected by both systems. However, Chandrasekar and Deming* have developed a procedure for estimation even in the presence of correlation.

This procedure is the division of events into sub-classes, each characterized by its own rate of omission. It can be shown that such a division leads to a more adequate estimate of the total number of events omitted from both register and survey. It is important to divide events into classes by a criterion that distinguishes frequent from infrequent omission. A promising and easily applied criterion for deaths is age at death; a useful two-way division is deaths under one year and all other deaths. Births may be separated into those of children recorded as dying before the first birthday, and all other. The basis for these suggested divisions is that infant deaths are usually reported in an especially incomplete manner, and births of non-surviving children are especially subject to omission.

Various elaborations of this basic experimental design can be imagined and might be profitable. Among the possibilities are the following:

1. Simultaneous experimentation with one or more techniques of registering and one or more household survey methods. If we designate the registration schemes as A and B and the enumeration schemes as 1 and 2, it would be possible to utilize interpenetrating subsamples using combinations A-1, A-2, B-1 and B-2. The result would be not only an estimate of birth and death rates in the area and of the completeness of the registration system being used in the non-sampled villages, but also an appraisal of the relative effectiveness of the alternative registry and survey schemes.

2. Additional sample villages could be added in which the survey system only and the registry system only were used in order to obtain an estimate of the effect on the completeness of the measurement of having a continuous check. The point is that the existence of matching will no doubt effect the performance of the registrar, of the enumerators, and possibly of the households themselves.

The advantages of this experiment, if it is at all successful, would be many. First, the result would be a reliable (and verified) estimate of vital rates. Second, the experiment would cumulate experience in the successful operation of a system of registration, which would make it possible to plan the gradual achievement of a state-wide or nation-wide registration of vital events. Third, the experiment would yield valuable data on the quality of survey information, and on the nature of recall lapse, and other systematic defects.

---

*Op. cit.

SOME RESULTS FROM ASIAN POPULATION STUDIES

William Seltzer

INTRODUCTION

The purpose of this paper is to bring together some results from a number of population growth estimation studies conducted in Asia since 1945 and to indicate a few generalizations suggested by these results. This review is limited to studies in which two collection systems are used simultaneously to gather basic data on vital events from the same population. Such a dual system provides a direct means by which the coverage of each source can be evaluated. Excluded are estimates based on a single, unevaluated census, household survey, or civil registration system, as well as estimates based on demographic analysis of census or survey data.

The studies considered are drawn from five countries ranging from Turkey in the West to Korea in the East and include data collected in normal civil registration systems, various special registration systems, and one-time and multiple-visit household surveys. Background data on the studies covered are summarized in Table 1.[1] This table provides information about the place where each study was conducted, the name of any larger study of which the listed study was a part, the coverage period (in terms of the date of occurrence of in-scope events), a brief description of each of the two collection systems used, and the approximate population covered. Unless otherwise mentioned in the description of the system, the recall period for multiple-visit household surveys extends to the date of the previous visit. In the case of sample studies the approximate population covered refers to the population in the sample.

Whenever possible, annual results of studies which used substantially the same methodology for a number of years have been combined into one set of coverage estimates for the full period for which data are available. This type of pooled experience is listed as a single study in Table 1. On the other hand, national sample studies for which coverage estimates can be derived for different geographical regions are listed separately in Table 1 for each such area. Basically, then, the 19 listings in Table 1 refer to a series of populations, in each of which reports of vital events occurring

[1] Omitted from Table 1 and this review are two studies in which the matching was done only 'one-way', i.e. the number of unmatched reports of only one of the two systems could be identified. For details, see Department of Census and Statistics, *Post Enumeration Survey 1953*, Monograph No. 1 (Ceylon Government Press, 1953), and Central Bureau of Statistics, Government of Indonesia, 'A Study of the Completeness of Birth and Death Statistics in Indonesia', Paper prepared for the ECAFE Seminar on Civil Registration and Vital Statistics, Copenhagen, July–August 1968. In addition, no comparable coverage estimates could be prepared on the basis of results currently available for the 1967 Ceylonese registration check, for a number of states participating in the Indian Sample Registration Scheme, and for the continuing study in the Lulliani area of West Pakistan conducted by the West Pakistan Research and Evaluation Centre, Lahore. Preliminary information available from the Sample Registration Scheme for states that are not covered in this review and from the Lulliani Project appear to be consistent with the general findings presented here. Undoubtedly other relevant studies have been omitted. However, their exclusion was inadvertent and the author would welcome assistance in locating them.

Reprinted from *Population Studies*, Vol. 23, No. 3, pp. 395-406 with permission of the publisher.

TABLE 1. *Background data on population growth estimation studies conducted in Asia*
(Four sources, see notes at end of table)

Study number	Country	Period covered	Study name and area within country	Description of first system	Description of second system	Approximate population covered
1	India	1945	Singur Health Centre, Bengal	Civil registration	One-time household survey with approx. two-year recall	64,000
2	India	1946	Singur Health Centre, Bengal	Civil registration	One-time household survey with approx. one-year recall	64,000
3	India	1950–52	Mysore Population Study – three rural zones	Civil registration	One-time household survey with recall period varying from about one to two years	23,300
4	India	1950–52	Mysore Population Study – Bangalore City	Pregnancy history in one-time fertility and attitude survey with female interviewer	One-time household survey with recall period varying from about one to two years	1,011
5	India	1964	Gujarat sample registration scheme (Pilot) rural	Special registration by local resident serving as part-time registrar	Household survey conducted every six months by supervisor	36,000
6	India	*Holi* 1964–May 1966	Vaso (Pilot) urban, Gujarat	Civil registration	One-time household survey with recall period of up to 28 months	9,480
7	India	December 1963–November 1966	Five villages in Mehrauli Block, South Delhi	Special registration by part-time non-resident registrar making fortnightly visits to local informants	Household survey conducted every six months by female interviewer	5,582
8	India	July 1965–June 1967	Kerala Sample Registration Scheme	Special registration by local resident serving as part-time registrar	Household survey conducted every six months by supervisor	260,000

(*Continued*)

COMMUNITY HEALTH LIBRARY AND INFORMATION CENTRE

TABLE 1 (*continued*)

Study number	Country	Period covered	Study name and area within country	Description of first system	Description of second system	Approximate population covered
9	Korea	October 1963–October 1965	Taeduk Gun Matching Project	Civil registration	One-time household survey with recall period of up to two years	100,000
10	Pakistan	January 1962–December 1965	Population Growth Estimation (PGE), East Pakistan	Special registration by full-time registrar living in the area and visiting local contacts on a regular basis	Household survey conducted every three months with one-year over-lapping recall periods	42,000
11	Pakistan	January 1962–December 1965	Population Growth Estimation (PGE), West Pakistan	Special registration by full-time registrar living in the area and visiting local contacts on a regular basis	Household survey conducted every three months with one-year over-lapping recall periods	42,000
12	Thailand	July 1964–June 1965	Survey of Population Change (SPC)	Civil registration	Household survey conducted every three months	172,000
13	Turkey	November 1965–October 1966	Turkish Demographic Survey (TDS) – Region 1 (Central Anatolian Plateau) urban	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	32,500
14	Turkey	October 1965–September 1966	Turkish Demographic Survey (TDS) – Region 1 (Central Anatolian Plateau) rural	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	38,100
15	Turkey	November 1965–October 1966	Turkish Demographic Survey (TDS) – Region 2 (Black Sea Coastal Provinces) urban	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	29,700

(*Continued*)

TABLE 1 (*continued*)

Study number	Country	Period covered	Study name and area within country	Description of first system	Description of second system	Approximate population covered
16	Turkey	October 1965–September 1966	Turkish Demographic Survey (TDS) – Region 2 (Black Sea Coastal Provinces) rural	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	37,600
17	Turkey	December 1965–November 1966	Turkish Demographic Survey (TDS) – Ankara City	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	24,900
18	Turkey	January 1966 – December 1966	Turkish Demographic Survey (TDS) – Istanbul City	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	15,600
19	Turkey	February 1966–January 1967	Turkish Demographic Survey (TDS) – Izmir City	Special registration by part-time local registrar making monthly household visits	Household survey conducted every six months by staff supervisor	13,400

SOURCES used in preparing Tables 1 through 5:

Study 1. C. Chandrasekaran and W. Edwards Deming, 'On a method of estimating birth and death rates and the extent of registration', *Journal of the American Statistical Association*, **44**, 245 (1949), Table 1.

Study 2. C. Chandrasekaran and W. Edwards Deming, *loc. cit.*, Table 1.

Study 3. United Nations, *The Mysore Population Study*, United Nations, Department of Economic and Social Affairs, Population Studies No. 34 (New York, 1961), Tables 16.2 and 16.5.

Study 4. United Nations, *op. cit.*, Tables 16.6 and 16.7.

Study 5. D. C. Mehta and M. H. Shah, *Report on Sample Registration Scheme (Pilot), Rural Gujarat*, Director of Health and Medical Services (Health) Gujarat (Ahmedabad, 1966), p. 9. Data refer to resident population only.

Study 6. D. C. Mehta, *Report on Under-Registration Study (Pilot), Urban Vaso Project Gujarat*, State Bureau of Health Intelligence, Directorate of Health and Medical Services (Health) Gujarat (Amedabad, 1967), pp. 7–8. Data refer to resident population only.

Study 7. D. V. R. Murty and P. K. Jain, 'Report on Pilot Sample Registration Scheme in Five Villages in Mehrauli Block, South Delhi: December 1, 1963–November 30, 1966', Central Family Planning Institute, New Delhi, February 1967, Table 2. Mimeographed report labelled 'Preliminary'. Data refer to *de jure* population and were obtained by summing the results presented separately for three 12-month periods.

(*Continued*)

TABLE 1 (*continued*)

Study 8. Sample Registration Kerala State, 'Report No. 1 for the Period 1965–66', the Bureau of Economics and Statistics, Trivandrum, November 1967, Table V; and Sample Registration Kerala State, 'Report No. 2 for the Period 1966–67', the Bureau of Economics and Statistics Trivandrum, August 1968, Table 6. Coverage estimates are derived from the summed unweighted sample data presented for individual 12-month periods in the cited tables. The text of 'Report No. 1 for the Period 1965–66' indicates (page 15) a Chandrasekaran–Deming estimate of 112 births and 32 deaths for the category missed by both systems during the 12-month period ending June 1966, the comparable estimates derived from the data shown in Table V are 124 births and 35 deaths. The source of this minor discrepancy is not known.

Study 9. E. H. Choe, *Problems and Adequacy of Vital Statistics in Korea*, The Population Studies Centre, Publication Series No. 1, Seoul National University, Seoul, April 1967, Table 4. Excludes registered events from outside survey area and survey events with *Hojuk* (permanent domicile) outside survey area or in a different *Myeon* (township). For approximate population covered, see E. H. Choe, 'Registration and Other Estimates of Vital Events in the Republic of Korea' (ASTAT/CRVS/CR7), Paper presented at the ECAFE Seminar on Civil Registration and Vital Statistics, Copenhagen, July–August 1968, Table 8.

Study 10. Farhat Yusuf, 'Population Growth Estimation: Studies in Methodology I – Matching of Vital Events', Research Report No. 67 Pakistan Institute of Development Economics, Karachi, April 1968, Tables 3A and 3C. Data are based on unweighted *de facto* reports for eight sample areas in East Pakistan where both collection systems were in operation and summed over four calendar years. For approximate population covered, see Population Growth Estimation, *Report of the Population Growth Estimation Experiment: Description and Some Results for 1962 and 1963*, Special Publication, Pakistan Institute of Development Economics, Karachi, 1968, Table II.

Study 11. Farhat Yusuf, *op. cit.*, Tables 3B and 3D. Data are based on unweighedt *de facto* reports for eight sample areas in West Pakistan where both collection systems were in operation and summed over four calendar years. For approximate population covered, see 'Population Growth Estimation', *op. cit.*, Table II.

Study 12. National Statistical Office, 'Report of the Survey of Population Change: 1964–67', Office of the Prime Minister, Government of Thailand, Bangkok, February 1969, Tables 4 and 5. Mimeographed report labelled 'to be used as provisional report until publication is completed'. Data refer to *de jure* population of the full sample and cover the first year of operation only. The estimated number of events shown in Tables 4 and 5 are obtained from the sum of individual estimates for population sub-groups.

Studies 13–19. John C. Rumford, Yasar Heperkan and Nuray Fincancioglu, 'The Principles and Preliminary Results of the Turkish Demographic Survey', *Public Health Reports*, Vol. 83, No. 17 (July 1968), pp. 573–582, Table 4; and Turkish Demographic Survey, *Vital Statistics from the Turkish Demographic Survey, 1965–66*, School of Public Health, Ministry of Health and Social Welfare, Republic of Turkey, Ankara, 1967, Appendix Tables 1, 11 and 26.

in a specified time period were collected by means of two relatively unvarying systems. Each of these 19 listings is treated as an independent and distinct study in this review, whether or not it is part of a larger study.

DERIVATION AND LIMITATIONS OF COVERAGE ESTIMATES

In each of the studies listed in Table 1, data on live births and deaths were obtained by two methods and the reports on births and deaths collected in each system were matched with those collected by means of the other system in order to determine which were reported by both and which by either system alone. With this information one can, subject to certain limitations, estimate the completeness of coverage of each system and the proportion of events missed by both systems.[2]

Specifically, the proportions of births or deaths reported by the first system that are also reported by the second system are used to estimate the coverage of the second system.[3] Stated more simply, the match rate of the first system is the coverage estimate of the second system and the match rate of the second system is the coverage estimate of the first system. In algebraic terms

$$\frac{C}{S_1} = \frac{S_2}{N} \qquad (1)$$

where

C = events (births or deaths) common to S_1 and S_2 determined by matching;
S_1 = the number of events (births or deaths) reported by one system;
S_2 = the number of events (births or deaths) reported by the other system; and
N = total number of births or deaths.

Transposing expression (1), we have a method of estimating the total number of events adjusted for both underregistration and underenumeration:

$$\hat{N} = \frac{S_1 S_2}{C} . \qquad (2)$$

In fact, this estimate, which was used for many years by the U.S. National Office of Vital Statistics to adjust United States birth registration data for underregistration,[4] is identical to the Chandrasekaran–Deming estimate.[5] It is therefore possible to derive Chandrasekaran–Deming estimates from any study for which the total number of reports (or the number of unmatched reports) of both systems and the match rate of either system (or the number of reports common to both systems) are available, whether or not such estimates were explicitly made in the original study.

[2] C. Chandrasekaran and W. Edwards Deming, 'On a method of estimating birth and death rates and the extent of registration', *Journal of American Statistical Association*, **44**, 245 (1949), pp. 101–115.

[3] S. Shapiro, 'Development of birth registration and birth statistics in the United States', *Population Studies*, **4**, 1 (June 1950), p. 98.

[4] See, for example, Robert D. Grove, *Studies in the Completeness of Birth Registration: Part I. Completeness of Birth Registration in the United States, December 1, 1939 to March 31, 1940*, U.S. Bureau of the Census, Vital Statistics – Special Reports, Vol. XVII, No. 18, 1943; and Sam Shapiro and Joseph Schachter, *Birth Registration Completeness in the United States and Geographic Areas, Part I: Data for Each State*, National Office of Vital Statistics, Vital Statistics – Special Reports, Vol. 39, No. 2, 1954.

[5] This can be seen by substituting $C+N_1$ for S_1 and $C+N_2$ for S_2 in expression (2) and expanding. The result can be recombined into the usual Chandrasekaran–Deming formula, i.e. $C+N_1+N_2+(N_1N_2/C)$.

In order that coverage estimates based on matching studies should be unbiased, the two systems employed in each study must be statistically independent, the net matching error should be zero (i.e. the number of false matches equal to the number of false non-matches) and all spurious (out-of-scope) reports must be eliminated from each system. In the present context, statistical independence means that the probability of one system catching an event is not related to the probability of the other system including it. As Chandrasekaran and Deming pointed out,[6] lack of independence can be largely offset by preparing separate estimates for groups of events which have relatively similar coverage rates and summing the individual estimates. In the few instances where this approach was attempted only minor gains were obtained,[7] indicating that either independence was not a problem or the groups used were not actually homogeneous with respect to coverage. With regard to spurious reports, most of the studies listed in Table 1 re-investigated the non-matched reports in the field in order to detect and eliminate those which did not refer to in-scope events. While the effects of matching errors and the lack of independence often, at least partially, cancel out,[8] the existence of these biases must be kept in mind when interpreting coverage estimates.

COVERAGE ESTIMATES FOR BIRTHS AND DEATHS

Estimates of completeness of coverage by collection system for each study are shown in Table 2 for births and for deaths. The coverage estimates are calculated from the data shown in the sources listed beneath Table 1. In a few instances, the estimates of coverage in Table 2 differ from the coverage estimates provided in the sources themselves. This situation arose either when the original study did not make use of expression (1), or its equivalent, in preparing coverage estimates or when the coverage estimate of the original study was prepared on the basis of separate estimates for population sub-groups assumed to have similar coverage rates. The variations in the coverage estimates shown in Table 2 reflect true differences in coverage as well as the sources of bias mentioned earlier. In most cases lack of independence can be expected to overestimate coverage, matching errors either to over or to underestimate coverage and spurious reports to underestimate coverage. Among these 19 studies, Table 2 provides 38 coverage estimates for births – two per study – which range from 28 to 96%. The median proportion covered for births among these 38 collection systems was 77%. For deaths, the coverage estimates range from 23 to 91% with a median of 69%.

It is clear from these data that the estimated coverage of live births is typically higher than that for deaths in Asian PGE studies. Based on the 38 collection systems taken together, the median coverage rate for births is about 11% higher than that for deaths. Similarly, the estimated coverage of births is higher than that for deaths in all but four of the 38 systems.

[6] Chandrasekaran and Deming, *loc. cit.*, pp. 106–107.

[7] Estimates based on the summation of sub-group estimates were prepared in Thailand's Survey of Population Change, Pakistan's Population Growth Estimation Experiment, and India's Singur Health Centre Study. Only in the last-named study was any meaningful difference observed between the results of the disaggregated and the global methods of estimation. Even in this case the difference was evident for only one of the variables tried. See Chandrasekaran and Deming, *loc. cit.*, pp. 110–111.

[8] William Seltzer and Arjun Adlakha, 'On the Effect of Errors in the Application of the Chandrasekaran–Deming Technique', Paper prepared for the Population Council Seminar on the Chandrasekaran–Deming Technique: Theory and Application, New York, 23rd January, 1969.

The ratios of the omission rate for births (i.e. one minus the coverage rate) to the omission rate for deaths vary considerably among the individual collection systems – but for only one out of the 38 systems is the estimated omission rate for deaths more than 2·4 times the rate for births. (The median ratio of the omission rates is about 1·3.) Despite the higher omission rates for deaths compared to those for births, the resulting estimates of natural increase tend to be too low. This will be the case as long as the ratio of births to deaths (or of the birth rate to the death rate) in the

TABLE 2. *Estimated coverage of births and deaths for each collection system in selected Asian PGE studies*

(Coverage of each system expressed as a percentage of the estimated total for each study)

Study number	System 1			System 2		
	Type of system	Estimated coverage		Type of system	Estimated coverage	
		Births	Deaths		Births	Deaths
1	Civil registration	52	48	One-time survey	53	32
2	Civil registration	60	51	One-time survey	67	51
3	Civil registration	43	44	One-time survey	93	90
4	One-time survey	96	88	One-time survey	92	88
5	Special registration	92	91	Multiple-visit survey	89	89
6	Civil registration	86	82	One-time survey	65	51
7	Special registration	89	74	Multiple-visit survey	91	81
8	Special registration	83	85	Multiple-visit survey	92	89
9	Civil registration	43	33	One-time survey	28	23
10	Special registration	83	84	Multiple-visit survey	75	61
11	Special registration	76	77	Multiple-visit survey	71	56
12	Civil registration	84	63	Multiple-visit survey	87	80
13	Special registration	67	61	Multiple-visit survey	86	82
14	Special registration	67	61	Multiple-visit survey	82	70
15	Special registration	70	62	Multiple-visit survey	83	75
16	Special registration	67	64	Multiple-visit survey	82	72
17	Special registration	62	50	Multiple-visit survey	66	50
18	Special registration	67	68	Multiple-visit survey	77	70
19	Special registration	66	58	Multiple-visit survey	86	72

NOTE: For sources and limitations of data, see Table 1 and text.

population under study is greater than the ratio of omission rate for deaths to the comparable omission rate for births.[9] Since the ratio of births to deaths among the countries of Asia averages about 2·6,[10] rates of natural increase will be underestimated whenever the omission rates for deaths are not greater than about 2·6 times those for births. The available evidence thus suggests that unadjusted registration, survey or census results in many Asian countries tend to underestimate the rate of population growth.

[9] *Ibid.* See also, United Nations, *Methods of Estimating Total Population for Current Dates*, Population Studies No. 10 (New York, 1952), p. 38, footnote 1.

[10] The ratio quoted was calculated from estimates presented in United Nations, *Population and Vital Statistics Report: Data Available as of 1 October 1968*, Statistical Papers Series A, Vol. XX, No. 4, for the sovereign countries of Asia.

COVERAGE ESTIMATES BY TYPE OF COLLECTION SYSTEM

Each of the 38 collection systems can be classified into one of four broad types: civil registration, special registration, one-time household surveys or multiple-visit household surveys. The median and range of coverage estimates by type of system are shown in Table 3. In addition to the general limitations of the method of analysis used in this review, the comparative data shown in Table 3 are subject to additional limitations due to the small number of examples available for each type of system. Nevertheless, the marked differences in median estimated coverage are suggestive. For both births and deaths the civil registration systems have the lowest median, then come, in the order of increasing median coverage, one-time household surveys, special registration systems, and multiple-visit household surveys. Median estimated coverage among the multiple-visit household surveys is about 45% higher than that for the civil registration systems for both births and deaths.

TABLE 3. *Median and range of estimated coverage of births and deaths by type of collection system, selected Asian PGE studies*

(Coverage of each system expressed as a percentage of the estimated total for each study)

Type of system	Number of systems	Estimated proportion of births covered		Estimated proportion of deaths covered	
		Median	Range	Median	Range
All types	*38*	*77*	*28–96*	*69*	*23–91*
Civil registration	6	56	43–86	49·5	33–82
Special registration	12	73	62–92	66	50–91
One-time household survey	7	67	28–96	51	23–90
Multiple-visit household survey	13	83	66–92	72	50–89

NOTE: For sources and limitations of data, see Table 1 and text.

Except for the one-time household surveys, median coverage rates for births are between 10 and 15% higher than those for deaths, regardless of the type of system. However, among the single-visit household surveys median coverage for births is 31% above that for deaths. In low-mortality countries it has been found that single-visit retrospective household inquiries are particularly unsuitable for obtaining a full count of deaths in a population.[11] These data suggest that the same may be the case in Asia.

Shifting from the medians to the ranges presented in Table 3, it is clear that coverage both for births and for deaths varies very widely within each of the four type of systems. Thus it would appear that the application of any particular system is at least as important as the choice of the system itself. This variability also argues against relying on predetermined factors inferred from experience in other countries, or even from other studies in the same country, to adjust any particular set of data for underregistration or underenumeration of vital events.

[11] U.S. National Health Survey, *Hospital Utilization in the Last Year of Life*, Health Statistics, Series D, No. 3, U.S. Department of Health, Education and Welfare, Washington, D.C., 1961.

Another way of looking at coverage is to examine the proportoin of the estimated total number of events that are covered by both systems, covered by the first system only, covered by the second system only, and missed by both systems, i.e. to distribute the estimated total into the four Chandrasekaran–Deming categories: C, N_1, N_2 and X.[12] This breakdown is presented in Table 4 for births and Table 5 for deaths for the 19 studies listed in Table 1. For these studies the proportion estimated as missed by both systems is relatively small – the median being 5% of the estimated total for births and 9% for deaths. The median proportion covered by both systems is 56% for births and 44% for deaths. The remaining events were caught by only one of the two systems used.

TABLE 4. *Percentage distribution of estimated total births, by coverage category, for selected Asian PGE studies*

Study number	Estimated total births	Percentage distribution				
		Estimated total	Caught by both systems	System 1 only	System 2 only	Missed by both systems
1	2,908	100	27	24	25	24
2	3,744	100	40	20	27	13
3	746	100	40	3	53	4
4	412	100	88	8	3	1
5	1,591	100	82	10	7	1
6	806	100	56	30	9	5
7	874	100	81	8	10	1
8	18,964	100	76	7	16	1
9	13,268	100	12	31	16	41
10	8,719	100	63	21	12	4
11	9,138	100	54	22	17	7
12	7,305	100	73	11	14	2
13	1,234	100	57	9	29	5
14	2,041	100	54	12	27	6
15	904	100	58	12	25	5
16	1,664	100	55	12	27	6
17	914	100	41	21	25	13
18	439	100	59	17	19	5
19	331	100	56	9	30	5

NOTE: For sources of data and their limitations, see Table 1 and text.

There are considerable variations in these proportions among the various studies. In point of fact, these relative distributions are directly related to the estimated coverage rates shown in Table 2. Recalling from expression (1) that, assuming statistical independence,

$$\frac{C}{S_1} = \frac{S_2}{\hat{N}}$$

it follows that the proportion of the estimated total caught by both systems is equal to:

$$\frac{C}{\hat{N}} = \frac{C}{S_1} \cdot \frac{C}{S_2} = \frac{S_2}{\hat{N}} \cdot \frac{S_1}{\hat{N}} \,. \qquad (3)$$

[12] Chandrasekaran and Deming, *loc. cit.*, p. 102.

Similarly, the proportion of the estimated total missed by both systems, i.e. Chandrasekaran and Deming's 'fourth category', is equal to:

$$\frac{X}{\hat{N}} = \left(1 - \frac{C}{S_1}\right)\left(1 - \frac{C}{S_2}\right) = \left(1 - \frac{S_2}{\hat{N}}\right)\left(1 - \frac{S_1}{\hat{N}}\right). \qquad (4)$$

It follows from this that the proportion of the estimated total caught by both systems must always be smaller than the smaller match (or coverage) rate, while the relative contribution of the fourth category is always less than the smaller non-match (or omission) rate.

TABLE 5. *Percentage distribution of estimated total deaths, by coverage category, for selected Asian PGE studies*

Study number	Estimated total deaths	Percentage distribution				
		Estimated total	Caught by both systems	System 1 only	System 2 only	Missed by both systems
1	2,234	100	16	33	17	35
2	1,696	100	26	25	25	24
3	328	100	40	5	50	6
4	56	100	77	11	11	1
5	618	100	81	10	8	1
6	359	100	42	40	9	9
7	310	100	60	14	21	5
8	5,302	100	76	9	14	2
9	2,553	100	8	26	15	52
10	3,152	100	52	32	10	6
11	3,088	100	44	34	13	10
12	1,867	100	50	12	30	7
13	417	100	50	11	32	7
14	782	100	42	19	27	12
15	284	100	46	16	28	10
16	552	100	46	18	26	10
17	262	100	25	25	25	25
18	164	100	48	20	23	10
19	130	100	42	16	30	12

NOTE: For sources of data and their limitations, see Table 1 and text.

DISCUSSION

Because the present review attempted to be as inclusive as possible, there is considerable variation in the quality of the studies used in the analysis. This variation has two main components: (1) variations in the type, quality and independence of the collection systems, and (2) variations in the matching and verification procedures. A more comprehensive analysis would attempt to control for variations in either or both of these components. At a minimum, it would be desirable to examine the extent to which coverage estimates for a particular type of system are merely an artifact of the type of system with which it was most frequently paired. Unfortunately, such an analysis will have to be deferred until results from a larger number of matching studies are available.

Another question of some importance is whether or not the multiple use of data from the same larger study (e.g., the Turkish Demographic Survey is represented seven times in this analysis, Pakistan's PGE Experiment twice) gives undue weight to the experience of a particular set of procedures. It is apparent from variations in coverage rates by area that some real differences exist within the large national studies in either the application or the effectiveness of the procedures used. However, additional independent studies are needed in order to eliminate this source of uncertainty definitively.

The estimates of completeness of coverage that serve as the statistical basis of the present review are derived solely from the results of matching the individual reports of two simultaneous collection systems. The approach used completely disregards any other evidence about coverage or quality, whether based on a simple judgment of the 'reasonableness' of the estimates of the vital rates or on the use of techniques of demographic analysis.[13] For example, two quite inadequate household surveys matched with each other would yield erroneously high estimates of coverage if both systems were missing the same kinds of events. The fact that the estimated coverage rates were too high could be established by determining from external evidence whether or not the estimates of vital rates were too low. However, a decision about the 'reasonableness' of birth and death rate estimates is to a large extent a function of the training and experience of the demographer making the determination. To a lesser extent the accuracy of the various analytical techniques is also dependent on the skill of the demographer using them.[14]

Since the present review includes studies from five countries and a larger number of areas, the time and specialized knowledge needed to make such independent assessments of each estimate would have been considerable. Moreover, in a number of Asian countries there is so much uncertainty about current levels of fertility and mortality and the supporting demographic data are subject to such large errors, that a number of demographers would question the correctness of any judgment made. Nevertheless, the overall coverage patterns indicated by the present limited analysis tend to support the findings of others[15] who did make use of various kinds of external data to assess quality.

While the limited, non-judgmental analysis used may be justifiable in this review, it is certainly not appropriate when evaluating the quality of the estimates produced by a particular PGE study or for a particular country. In that situation one would want to make use of as many different sources of data and as many techniques of estimation as resources permit, since no one approach is free from error. Viewed in this light, estimates based on matching studies and those based on methods such as the Brass technique are complementary rather than competitive. It would appear that, unless only very rough approximations of current fertility and mortality are required, designs for new population growth estimation studies should provide both for a full matching operation and for the collection and tabulation of information needed to make use of analytical techniques.

[13] Ansley J. Coale and Paul Demeny, *Methods of Estimating Basic Demographic Measures from Incomplete Data*, United Nations, Department of Economic and Social Affairs, Population Studies No. 42 (New York, 1967).

[14] *Ibid.*, pp. 51–52.

[15] W. Parker Mauldin, 'Estimating rates of population growth', in *Family Planning and Population Programs*, edited by Berelson *et al.*, (Chicago, 1966), pp. 635–653; and Sultan S. Hashmi, 'Use of household surveys for measuring population changes in developing countries with special reference to Pakistan', *Proceedings of the CENTO Symposium on Household Surveys, Dacca, 1966* (Ankara, 1967), pp. 165–182.

Publications of the POPLAB Program:

The Scientific Report Series

1. THE CONCEPT AND THE PROGRAM OF THE LABORATORIES FOR POPULATION STATISTICS by Forrest E. Linder
2. ORGANIZATION AND METHODS OF THE DUAL-REPORT SYSTEM IN TURKEY by Eliska Chanlett

The Manual Series

1. MAPPING AND HOUSE NUMBERING by Dorothy S. Cooke

The Reprint Series

1. ON A METHOD OF ESTIMATING BIRTH AND DEATH RATES AND THE EXTENT OF REGISTRATION by C. Chandra Sekar and W. Edwards Deming
2. THE USE OF SAMPLING FOR VITAL REGISTRATION AND VITAL STATISTICS by Philip M. Hauser
3. THE DESIGN OF AN EXPERIMENTAL PROCEDURE FOR OBTAINING ACCURATE VITAL STATISTICS by Ansley J. Coale

---

SOME RESULTS FROM ASIAN POPULATION GROWTH STUDIES by William Seltzer